A Study on the Autoregressive Model and Forecasting for the Birth Rate in India

P. Rajalakshmi
Research Scholar
&
Dr. Lilly George
Assistant Professor, Department of Statistics
St. Joseph's College, Tiruchirappalli

Abstract

In this study, the main objective is to predictive in autoregressive model and forecasting for the birth rate in India. To predicts the nature of birth rate for the next few years. The value for Birth rate, crude (per 1,000 people) in India was 19.66 as of 2015, over the past 55 years this indicator reached a maximum value of 42.11 in 1960, and a minimum value of 19.66 in 2015. Further the analysis for the Predictive value is birth rate value of 19.39 in 2016and 2017 value of birth rate 19.14. Data collected: The data is secondary data collected from the website: http://www.intexmundicom/facts/India/birh rate. The data is collected yearly wise Birth rate from 1960 to 2015.

Keywords: Statistical tools, Forecasting with auto regressive model, Statistical packages for social science

1. Introduction

Definition: Crude birth rate indicates the number of live births occurring during the year, per 1,000 population estimated at midyear. Subtracting the crude death rate from the crude birth rate provides the rate of natural increase, which is equal to the rate of population change in the absence of migration.

Basic Concepts

Crude Birth Rate: The number of live births per year per thousand of the population.

$$\text{CRD} = \frac{\text{Total births in a specific year}}{\text{Total population in that year}} \times 1000$$

Crude birth rates are very general statistics that are useful for making overall comparisons between different groups of peoples, states, societies and regions.

Autoregressive Model

An autoregressive model is when a value from a time series is regressed on previous values from that same time series data. for example,

$$\hat{Y}_t = \beta 0 + \beta 1 Y_{i-1} + \epsilon t \qquad ...(1)$$

In this regression model, the response variable in the previous time period has become the predictor and the errors have our usual assumptions about errors in a simple linear regression model. The order of an autoregression is the number of immediately preceding values in the series that are used to predict the value at the present time. So, the preceding model is a first-order autoregression, written as AR(1).

If we want to predict y this year (yt) using birth rate datae in the previous two years (Y_{i-1}, Y_{i-2}) then the autoregressive model for doing so would be:

$$\hat{Y}_t = \beta 0 + \beta 1 Y_{i-1} + \beta 2\, Y_{i-2} + \epsilon t \qquad ...(2)$$

This model is a second-order autoregression, written as AR(2), since the value at time t is predicted from the values at times t−1 and t−2. More generally, a Kth-order autoregression, written as AR(k), is a multiple linear regression in which the value of the series at any time t is a (linear) function of the values at times t−1,t−2,…,t−k.

2. Review of Literature

Fertility indicates the reproductive performance of a woman. The child bearing period of a woman is generally assumed to exist between the age 15 to 40. The level of fertility in demography is measured in terms of live-birth performance. The child bearing is, no doubt, basically a biological function, but the child bearing in any society is performed in socio-economic and cultural setup and is, therefore, influenced by socioeconomic factors as well as social customs, values and norms related to various aspects of childbearing Bhende (2003).

Researches conducted in the field of fertility relate to one or more of the stratification variables like income, education, family type and occupation. These stratification variables according to the Freedman (1963) are important in fertility research because of their role in social and economic development. Many studies were made to explain fertility differentials in evaluating the prospects for an early fertility decline in India (Parkasi and Malakar, 1967; Rele, 1972; Misra and Sharma, 1978; Reddy, 1978). Driver (1960) reported that there were slight variations in the mean number of children born in different income groups amongst Muslims, Hindus, Buddhists and commented that education and fertility showed inverse relationship with the fertility i.e. with increase in education the fertility level declined amongst these caste groups in India. An illiterate wife and husband had 4.7 and 5.0 children born respectively, whereas wife of above primary level education and father with college education showed 3.9 children each. While studying the impact of occupation on fertility, he found that unskilled workers (4.3), agriculturists (4.9) and artisans (4.5) had the highest fertility, whereas the clerks had the lowest (3.9).

The joint survey of Mysore population (1961) by United Nations and Govt. of India revealed inverse relationship between education and fertility, that is, women with high school or university education had 4 and 2 children as compared to 5.4

for those with less education. The study also found little difference in mean number of children born to illiterate and literate women in Bangalore city, however, lower fertility was observed among educated women. The type of dwelling was used as an index of economic status by them, and they observed a positive association between completed fertility (that of women 45 years of age or over) and upper status. Female who were living in huts had 4.4 live-births, whereas those living in mud-houses and in still better houses had 4.5 and 5 live births respectively.

Anand (1966) observed a positive correlation between the education of the woman and number of children born/pregnancies; number of children was highest (9.1) among illiterates whereas the graduate mothers had 7.0 pregnancies. Illiterate as well as primary educated fathers showed 7.7 children, graduate and that with technical or professional education had 7.2 and 6.2 children respectively. The study also indicated inverse relationship between income and fertility, with more number of children in lowest income group as well as lower number of children amongst higher income group people. The number of children declined with an increase in the income except in Rs. 600 to 700 group, where it declined again, thereby showing non-significant correlation between income and fertility.

Saxena (1973), in his survey of Lucknow city showed a clear inverse relationship between general marital fertility rate and monthly income. The general marital fertility rate was 240 for persons having income less than Rs. 75 per month, 180, 140 and 80 for people belonging to income categories of Rs. 150-300, Rs.500-750 and Rs. 750 and over respectively. While studying fertility, he found that birth rate of illiterate women was 40.2 and that of educated was 36.7. He found significant negative association between fertility rates and education.

George (1976) found inverse relationship of income and occupation with fertility. He found high rank officials had lower fertility than other occupations viz. semi-skilled and skilled workers at all age groups, however, respondents from higher income group had smaller families than lower income groups in all age groups. This study showed strong inverse relationship between education and the number of live-births, illiterates had more children (5.57) and those with high school and college education had 4 and 3.87 children respectively.

Mohammed (1977) studied the effect of socioeconomic variables on fertility in the Economic and Social variables on fertility in the Economic and Social Commission for Asia and the Pacific (ESCAP) region and found that the number of physician, literacy and economic activity rates inversely related to fertility.

Nair (1978) found higher female literacy rates, medical facilities and favourable policy of government towards family planning programme responsible for bringing down the birth rate. Reddy (1978) studied the differential fertility in India and observed lower fertility among women with high school education than the literate women. He also found negative association between education

attainment and the fertility level. Sharma and Mishra (1978) observed a negative association between education and marital fertility and a decrease in fertility with rising educational status of the mother. The total marital fertility was found to be highest (5.29) for illiterate mother than that of mothers with higher education (3.71) among different caste groups in India. The investigators concluded that increasing literacy and educational attainment, higher age at marriage and acceptance of family planning would induce the fertility decline.

Mahedevan (1979) in his investigations on determinants of fertility differentials in South India, studied the variation in fertility with the help of socio-cultural variables, viz., age and family income, education and occupation of husband etc. Family income was found to be positively associated with the fertility. The study showed inconsistent pattern of fertility difference by occupation. For the population as a whole, fertility was found to be highest among agricultural laborers and lowest among owners and cultivators. As the general literacy level of wives was uniformly low among all the groups, therefore, the literacy rate of the husband was considered for educational variations in fertility differences. Education did not differ much as the proportion of higher educated men was very small among all the groups. Education of the husband was inversely related to the fertility in this study.

Singhal (1980) observed in Tripura that the level of education of the women tend to reduce the general marital fertility ratio. His conclusion was that with social and economic development, increasing literacy and higher marriage age, there would be continuous reduction in fertility in future.

Varadarajan (1981) surveyed the Kotas in Niligiris district (Tamil Nadu) and found that women belonging to the house-holds having annual income of Rs. 2000 and below, Rs. 2001-4000, Rs. 4001-6000 and Rs. 6001-8000 and Rs. 8001 and above had 2.74, 4.05, 3.06, 3.15 and 2.75 average number of live-births respectively. The study indicated the highest average of live-births for illiterate women (4.16), followed by those who had primary (2.54) and secondary education (2.29). Women had the highest live-births (4.18), if their husbands were illiterate, while those husbands had primary, secondary and college education had 3.57, 2.78 and 1.60 live-births respectively, thereby showing inverse relationship between education and fertility.

Krishnamoorthy (1980) noted that a desire for continuity of the family name or gender preference and for living children to support their old age is thought to influence the fertility behaviour of couples in many countries. A reduction in mortality will increase the chances of survival for children and the continuation of family name. Thus, it is argued that with improved conditions, couples would tend to limit their reproduction, since a small number of births would be sufficient to assure continuity of the family name and survival of one or two children to support the parents during their old age.

Bhuyan and Ahmed (1984) illustrated inverse but weak relationship between educated husband and fertility, with an appreciable decline of fertility at graduate level. An inverse relationship between the educational attainment of the mother and fertility was seen in this study, female education had more depressing effect upon fertility than male education. The study further revealed that occupation of husbands had significant impact on fertility. Persons involved in agriculture had maximum number of children born (4.36) followed by labourer, business, service and other technicians etc. This differential in fertility among husbands with different occupational patterns was highly significant.

Patnaik (1985) demonstrated the socio-economic and demographic determinants of fertility behaviour in Patna district. The findings of the study indicated that fertility of population varied by their socio-economic and demographic characteristics. Education was also an important factor influencing the fertility. The present study indicated an inverse relationship between education and fertility. It had been observed that as the level of education of husband changes from category of illiterates to the categories of literate, matriculate, intermediate, graduate and post-graduate, the decline in fertility had been of the order of 16, 27, 39, 43 and 44 percent respectively. Similarly, the respective declines in mean fertilities, when couples were classified by educational level of wives, were of the order of 30, 34, 42, 47 and 54 per cent.

Mathur (1986) reported an association between education of the husband and income level up to some extent. The study further indicated a positive relationship between monthly income of respondent males of different occupations with number of children born to their wives, belief in family planning, knowledge as well as use of family planning methods, opinion on government facilities in regard to family planning and their attitude towards inclusion of sex.

Puri (1989) illustrated that fertility level declined sharply as the educational level of the women increased, while the total fertility rates for illiterate women was 5.1, it was 2.1 only for women, who were literate but below middle school level. Thus, he observed that fertility levels were declined as educational level increases. Haile (1990) examined the socio-cultural, economic and demographic characteristics of 734 women aged 15-55 in the Gondar administrative region of North-Western Ehiopia and found that women had few children than desired and stopped child bearing when they reached or closely approximated their ideal number of sons. Since son was clearly, the determinant of reproductive success, it is argued that only a significant change in the status of women can bring about widespread compliance with the official family planning's two child norm.

Lloyd (1991) found a consistent negative association between women's paid work and fertility in developing countries. He reported that effect of work on fertility in the short run is contrary to its effect in the long run. In the long run women who have ever worked, end up with fewer children, but in the short-run, current work appears to be associated with higher fertility. In almost all the countries, differences between women in their occupational experience (work-

status) were found to be statistically significant in accounting for differences between women in their number children ever-born.

Das and Pandhiyar (1991) reported that the levels of total marital fertility rate have a tendency to decrease with an increase in the educational level of the husband and wife. Similarly, family income was found to be negatively related with fertility. The study indicated that the fertility of manual workers was higher as compared to that of white-collar workers. The use of conception has a tendency to increase with an increase in socio-economic status, while the duration of postpartum in fecund ability has a tendency to decrease. It was found that husband's occupation and family income were also important in causing favourable changes in fertility by off-setting the fertility-enhancing effect of modernization. However, the education of the husband and wife indicated little effect in lowering fertility.

Murthi et al. (1995) illustrated that female literacy and female labour force participation had a negative and statistically significant effect on TFR. Fertility was also significantly lower in the southern and western regions.

Syamala (2001) demonstrated the influence of childhood mortality on fertility behaviour of women in Goa. They showed that the net effect of child mortality could be substantial. Women with personal experience of child loss and having pessimistic opinion about the level of mortality, produced, on an average, about two children more than similar women who never experienced a child loss and were optimistic about the level. The tendency to replace a dead child was found to cut across the level of literacy and religious background of women and hence increased fertility rate.

Awah and Zuberi (2001) examined the association between childhood mortality and socioeconomic status (housing characteristics and household possession such as source of water, type of toilet facilities, housing construction materials; and household possession like radio, television, and animal possessions) in three southern African countries. They found that the chances of childhood mortality decreased consistently with levels of the socioeconomic status index.

Letamo and Oucho (2002) conducted a study on contribution of family planning programmes to fertility decline and found that the contribution of family planning methods and the use of modern contraceptives had a greater impact on fertility reduction in Botswana. They reported that family planning practices should also be seen as joint responsibility of both men and women. It was also evident from the study that men do influence the uptake of contraception and as such educating men to appreciate the importance of family planning was likely to lead to increase use of contraception which helped in reduction of fertility.

Singh *et al*. (2002) studied the impact of education and autonomy on fertility of women of rural and semi- urban areas of Varanasi district of Eastern Utter Pradesh of India. They found that level of autonomy increases as education increases. The percentage of women possessing high level of autonomy among

highly educated, was twice of that relating to uneducated. They concluded that substantial reduction in fertility can be achieved through popularizing women education, promoting of employment opportunities for women, improving in women's role in decision making and encouraging inter-spousal communication in family affairs. The desire for son was associated with low contraceptive prevalence. It had been suggested that the value of sons would be lowered in order to make family planning a success in reducing fertility.

Breierova and Duflo (2003) estimated the effect of education on fertility and child mortality by a school construction program that took place in Indonesia between 1973 to 1978. They showed that female education is a strong determinant of age at marriage and early fertility than male education. However, male and female education seems equally important factors in reducing child mortality. Wardle and Steptoe (2003) investigated attitudes and beliefs that might underlie behavioural choices, including health locus of control, future salience, subjective life expectancy, and health consciousness, in a nationally representative sample. Lower socioeconomic status was associated with less health consciousness, stronger beliefs in the influence of chance on health, less thinking about the future, and lower life expectancies. Socioeconomic differences in healthy lifestyles were associated with differences in attitudes to health that may themselves arise through variations in life opportunities and exposure to material hardship and ill health over the life course.

Biswas and Kapoor (2004) studied the age at menarche and menopause of the women among Saharia – a primitive tribe of Madhya Pradesh. In women life, menarche and menopause were the significant and inevitable events which indicated a particular adult stage of first periodical flow of blood from womb and cease to ovulate and menstruate respectively in all healthy women. Mean age at menarche and menopause of women were varied in regarding the different background factors like – nutritional status, family size, medical facility, genetic factor, environment, education, birth rank, living standard, socio-economic condition, etc. Mean menarcheal and menopausal age of Saharia women have been found to be high (13.5 ± 0.84 years) and low 44.6 ± 1.17 years) respectively.

Nahar and Rahman (2006) examined the factors associated with women's age at first marriage and interval between marriage and birth during 1983-1985 and 1992-1994. They documented that age at marriage was steadily rising in rural areas of Bangladesh, and the increase was strongly associated with increased female education. However, education of women has opposite effects on childbearing i.e. on the first birth (fertility). There was an indication that educated women tend to have their first birth as early as possible after marriage.

Josipovic (2007) examined the relationship between the education and fertility. In the prevailing view of the literature, education was treated as a highly influential factor in reproductive behaviour. This view posited a simple linear relationship between the two. It was argued that education and fertility make, rather than a simply linear relationship, a subtle pair. The data analyzed was

selected from geographical regions of Slovenia. The analysis brought curved relationship between education and fertility. He concluded that disparities in fertility rates among women with dissimilar levels of education in the studied territories were slowly diminishing through time, or have converted into other kinds of factor relations.

Hagestad and Call (2007) documented that overall differentials in completed fertility can be assessed both in terms of having any children and number of children. In low-fertility countries, a negative association between educational level and both having any children and number of children was often found in women, although lately differences in Scandinavian countries had been found to be relatively modest. For men there was less evidence on this issue and it was also less consistent than for women.

Goodman and Koupil (2009) Studied biological and social determinants of mortality and fertility which provided insight into selective pressures in a population. They studied determinants of reproductive success using multi-generational data from a large, population-based cohort of 13,666 individuals born in Sweden between 1915 and 1929. The effects of birth order, mother's age, mother's marital status and family socioeconomic position (SEP) upon reproductive success, measured as total number of children. Reproductive success was associated with both social and biological characteristics at birth. Higher family SEP was also associated with improved fertility.

Priya *et al*. (2009) reported that menarche and menopause demarcated the limits of potential reproductive life span in the female. A number of studies had conducted on this aspect in different endogamous population of Andhra Pradesh at different time’s. For the present work the urban and rural areas of Kshatriya women had been taken to study and this population has not studied previously on this aspect. In the study population it had been observed that, the early menarche in urban area while late menarche more in rural area. The difference of mean menopausal age between rural and urban areas was not significant.

Huber *et al*. (2010) investigated the association between socioeconomic status and reproductive output which was varied by the source of status and resources, the woman's education, and her age at reproductive onset (proxied by age at marriage). They examined the association between a woman's reproductive output and income and education by using a large sample of US women. Education, income, and age at marriage were negatively associated with a woman's number of children and increased her chances of childlessness. Among the most highly educated two-thirds of the sample of women, husband's income predicted the number of children. The association between a woman's number of children and her husband's income turned from positive to negative when her education and age at marriage was low (even though her mean offspring number raised at the same time). The association between a woman's own income and her number of children was negative, regardless of education.

3. Statistical analysis

Table 1
Developing first & second-order, autoregressive models on annual crude birth rate (1960-2015)

Year	Birth Rate	Lag1	Lag2	Year	Birth Rate	Lag1	Lag2
1960	42.11			1988	32.69	33.28	33.84
1961	41.87	42.11		1989	32.10	32.69	33.28
1962	41.60	41.87	42.11	1990	31.50	32.10	32.69
1963	41.32	41.60	41.87	1991	30.91	31.50	32.10
1964	41.01	41.32	41.60	1992	30.34	30.91	31.50
1965	40.70	41.01	41.32	1993	29.79	30.34	30.91
1966	40.37	40.70	41.01	1994	29.27	29.79	30.34
1967	40.06	40.37	40.70	1995	28.77	29.27	29.79
1968	39.75	40.06	40.37	1996	28.29	28.77	29.27
1969	39.45	39.75	40.06	1997	27.82	28.29	28.77
1970	39.15	39.45	39.75	1998	27.37	27.82	28.29
1971	38.85	39.15	39.45	1999	26.91	27.37	27.82
1972	38.53	38.85	39.15	2000	26.46	26.91	27.37
1973	38.19	38.53	38.85	2001	26.00	26.46	26.91
1974	37.84	38.19	38.53	2002	25.55	26.00	26.46
1975	37.49	37.84	38.19	2003	25.09	25.55	26.00
1976	37.16	37.49	37.84	2004	24.62	25.09	25.55
1977	36.85	37.16	37.49	2005	24.15	24.62	25.09
1978	36.58	36.85	37.16	2006	23.65	24.15	24.62
1979	36.33	36.58	36.85	2007	23.14	23.65	24.15
1980	36.10	36.33	36.58	2008	22.63	23.14	23.65
1981	35.85	36.10	36.33	2009	22.10	22.63	23.14
1982	35.57	35.85	36.10	2010	21.60	22.10	22.63
1983	35.23	35.57	35.85	2011	21.12	21.60	22.10
1984	34.83	35.23	35.57	2012	20.68	21.12	21.60
1985	34.37	34.83	35.23	2013	20.29	20.68	21.12
1986	33.84	34.37	34.83	2014	19.95	20.29	20.68
1987	33.28	33.84	34.37	2015	19.66	19.95	20.29

Table 2
Model Summary

Model	R	R Square	Adjusted R Square	Standard Error of the Estimate
1	1.000[a]	1.000	1.000	.02618

a. Predictors: (Constant), lag2, lag1

Conclusion: The coefficient of determination is 1.000; therefore, about 100% of the variation in this data is explained by birth rate. The regression equation appears to be very useful for making predictions since the value of $r^2$is close to 1.

Table 3
ANOVA[b]

Model		Sum of Squares	Df	Mean Square	F	Sig.
1.	Regression	2482.686	2	1241.343	1.811E6	.000[a]
	Residual	.035	51	.001		
	Total	2482.721	53			

a. Predictors: (Constant), lag2, lag1
b. Dependent Variable: birth_rate

Conclusion: This indicates the statistical significance of the regression model that was run. Here, $p < 0.001$, which is less than 0.05, and indicates that, overall, the regression model statistically significantly predicts the outcome variable (i.e., it is a good fit for the data).

Table 4
Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.071	.036		1.982	.053
	lag1	2.028	.042	2.010	48.187	.000
	lag2	-1.030	.043	-1.010	-24.221	.000

a. Dependent Variable: birth_rate

The fitted model is $\hat{Y}_i = .071 + 2.028\, Y_{i-1} - 1.030\, Y_{i-2}$ ...(3)

Here we take 2nd order coefficient and test for significance.

The hypotheses used for this test are,

$H_0 : A2 = 0$

$H_1 : A2 \neq 0$

Conclusion:

The p-value of $0.001 \leq \alpha = 0.05$, reject H_o. conclude that the second order parameter of the autoregressive model is significant. The model build approach has led to the selection of the second order autoregressive model as the most appropriate for the given data.

Using the estimates ao=0.071,a1=2.028and a2=-1.030 as well as two resent data value is Y_{55}=19.95 Y_{56}=19.66,the forecasts of birth rate for the year 2016,and 2017 are obtained from equation.

$$\hat{Y}_i = .071 + 2.028\ Y_{i-1} - 0.030\ Y_{i-2} \quad ...(4)$$

2016: 1 year ahead $Y_{55} = 0.071+2.028(19.66)-1.030(19.95) = 19.39$

2017: 2 year ahead $Y_{56} = 0.071+2.028(19.39)-1.030(19.66) = 19.14$

2018:3 year ahead $Y_{57} = 0.071+2.028(19.14)-1.030(19.39) = 18.95$

4. Findings & Conclusion

The p-value of $0.001 \leq \alpha = 0.05$, reject H_o. conclude that the second order parameter of the autoregressive model is significant. The model build approach has led to the selection of the second order autoregressive model as the most appropriate for the given data.

Using the estimated value for the birth rate in 2016 is 19.39 and 2017 is 19.14 and 2018 is 18.95, so declining the birth rate year by year.

Some important reasons for declining rate of crude birth rate are:

- increased focus on family planning,
- increase in wages,
- reduced subsistence agriculture and
- increase in education level of the women.

Government and non-governmental organizations have been encouraging family planning methods that are cheap and manageable. More than 50% of people living in the urban areas have accepted the use of family planning methods leading to low birth rates. Working people in urban areas want better pay, implying that they have to reduce the number of children so as to increase the time they spend at their workplace. The subsistence agriculture has been decreasing in the recent years. This has led to the rise in the cost of living therefore making it impossible to manage a big family. Women in the recent years are spending more time increasing in education for high paying jobs.

Refrence

1. Damodar N. Gujarati. “Basic econometric”, 2009.
2. David M. Levine, Timothy C. Krehbiel, Mark L. Berenson, “Business Statistics - A first course”, third edition, 2007, pp. 591-599.

3. Naval Bajpal, “Business Statistics”, 2010, pp. 613-615.
4. http://www.intexmundicom/facts/India/birh rate
5. Joseph, "Chaos Forecasting Insights," *Future Trends Newsletter*, Vol. 24, No. 2, (1993), p. 1.
6. Joseph, II, "Quality Approaches to Long-Range Forecasts," *Futurics: A Quarterly Journal of Futures Research*, Vol. 16, Nos. 3 & 4, (1992), p. 14.
7. Makridakis, "The Art and Science of Forecasting," *International Journal of Forecasting*, Vol. 2 (1986), p. 45.
8. Pohl, "The Uses of the Future," *The Futurist*, March-April, (1993), p. 9.
9. George, "Forecasting, Planning and Strategy for the 21st Century", *Futurics: A Quarterly Journal of Futures Research*, Vol. 16, Nos. 3 & 4, (1992), p. 56.